



Viral Communities in the Global Deep Ocean Conveyor Belt Assessed by Targeted Viromics

Daniele De Corte^{1*}, Joaquín Martínez Martínez², Mariana Silvia Cretoiu², Yoshihiro Takaki³, Takuro Nunoura⁴, Eva Sintes^{5,6}, Gerhard J. Herndl^{5,7} and Taichi Yokokawa³

¹ Research and Development Center for Marine Biosciences, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan, ² Bigelow Laboratory for Ocean Sciences, Boothbay, ME, United States, ³ Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan, ⁴ Research Center for Bioscience and Nanoscience (CeBN), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan, ⁵ Department of Limnology and Oceanography, Center of Functional Ecology, University of Vienna, Vienna, Austria, ⁶ Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Palma, Spain, ⁷ Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research, Utrecht University, Utrecht, Netherlands

OPEN ACCESS

Edited by:

Rui Zhang,
Xiamen University, China

Reviewed by:

Simon Roux,
Joint Genome Institute, United States
Marla Trindade,
University of the Western Cape,
South Africa

*Correspondence:

Daniele De Corte
daniele.de.corte@jamstec.go.jp

Specialty section:

This article was submitted to
Aquatic Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 20 February 2019

Accepted: 22 July 2019

Published: 21 August 2019

Citation:

De Corte D, Martínez Martínez J, Cretoiu MS, Takaki Y, Nunoura T, Sintes E, Herndl GJ and Yokokawa T (2019) Viral Communities in the Global Deep Ocean Conveyor Belt Assessed by Targeted Viromics. *Front. Microbiol.* 10:1801. doi: 10.3389/fmicb.2019.01801

Viruses are an abundant, diverse and dynamic component of marine and terrestrial ecosystems. In the ocean, viruses play a key role in the biogeochemical cycles and controlling microbial abundance, diversity and evolution. Recent metagenomics studies assessed the structure of the viral community in the upper ocean. However, little is known about the compositional changes in viral communities along the deep ocean conveyor belt. To assess potential changes in the viral community in the global deep-water circulation system, water samples were collected in the core of the North Atlantic Deep Water (NADW) (~2,500 m) and Pacific Antarctic Bottom Water (~4,000 m). Microbial and viral abundance were evaluated by flow cytometry. Subsequently, flow cytometry was used to sort virus-like particles and next generation sequencing was applied to build DNA libraries from the sorted virus populations. The viral communities were highly diverse across different oceanic regions with high dissimilarity between samples. Only 18% of the viral protein clusters were shared between the NADW and the Pacific Antarctic Bottom Water. Few viral groups, mainly associated with uncultured environmental and uncultured Mediterranean viruses were ubiquitously distributed along the global deep-water circulation system. Thus, our results point to a few groups of widely distributed abundant viruses in addition to the presence of rare and diverse types of viruses at a local scale.

Keywords: targeted viromics, deep ocean circulation, viruses, deep ocean, next generation sequencing

INTRODUCTION

Viruses shape marine ecosystems by controlling their host abundance and diversity through cell lysis and generate and maintain diversity through horizontal gene transfer (Angly et al., 2006; Rohwer and Thurber, 2009; Zeng and Chisholm, 2012). Moreover, viral lysis influences oceanic productivity by promoting organic matter and nutrient cycling through the release of intracellular material from the host cells (Middelboe et al., 1996; Middelboe and Lyck, 2002).

Several parameters affect the distribution of viruses in the marine environment (De Corte et al., 2012, 2016; Brum et al., 2015). The presence of a suitable host is key to perpetuate the viral

“life cycle.” However, other variables, such as temperature (Parada et al., 2007), UV radiation (Wommack et al., 1996) and salinity (Kukkaro and Bamford, 2009), influence the time span that viruses remain active in the ambient water and capable of infecting a host. The effects of these biotic and abiotic factors have been studied in surface waters (Bongiorni et al., 2005; De Corte et al., 2011) and, infrequently, in the deep ocean, where low temperature leads to relatively slower virus decay rates than in surface waters (Parada et al., 2007). The longer bacterial generation time and slower viral decay rates in the deep ocean (Parada et al., 2007; De Corte et al., 2012), together with the limited vertical mixing of the deep-water masses, suggest that oceanic currents and the thermohaline circulation may play an important role in influencing the global distribution of viruses and their hosts (McGillicuddy et al., 2007; Sintes et al., 2013, 2015; Sul et al., 2013; Brum et al., 2015; Frank et al., 2016).

The global overturning circulation (Figure 1) is a system of oceanic currents driven by wind, density and mixing processes (Schmittner et al., 2007; Talley, 2013). The North Atlantic Deep Water (NADW) is formed in the North Atlantic Ocean through the sinking of cold and dense waters, and moves southward toward the Southern Ocean (Talley, 2013). The second limb of the deep overturning circulation system is the Antarctic Bottom Water (AABW). The AABW is formed in the Southern Ocean via mixing of NADW, Indian Deep Water, Pacific Deep Water, and shelf waters originating from Antarctic shelf systems (Talley, 2013).

Viruses, as part of the dissolved organic matter pool (DOM; operationally defined as the organic matter that passes through a 0.2 μm filter), should behave conservatively and reflect the different formation sites of the deep waters (Bercovici and Hansell, 2016). However, it is important to note that certain local effects, such as particle export and *in situ* production (Nagata et al., 2000; Hansell and Ducklow, 2003; Yokokawa et al., 2013), play an important role in determining bacterial and viral abundance and community composition. Sinking particles provide a source of fresh substrate for microbes in the deep sea (Cho and Azam, 1988; Karl et al., 1988; Follett et al., 2014). Hence, this organic matter might be especially important in the Pacific Deep Water of the North Pacific with its low DOM concentration (Hansell, 2013).

DNA viruses lack conserved marker genes, hindering the assessment of their biogeographical patterns. The dispersal of viruses across the oceans has only been recently explored (Angly et al., 2006; Hurwitz and Sullivan, 2013; Martinez-Hernandez et al., 2017; Gregory et al., 2019). Moreover, most studies addressing viral community composition are limited to the upper ocean layers (Angly et al., 2006; Brum et al., 2015; Hurwitz et al., 2015). Due to methodological limitations, virus-host interactions in the dark ocean have mainly been assessed by microbial and viral enumeration (De Corte et al., 2012; Wigington et al., 2016; Lara et al., 2017). Therefore, little is known about the changes of viral communities along the deep ocean conveyor belt circulation system.

Although viromics (metagenomics of viral communities) is increasingly used to study viral diversity and function in the marine ecosystems (Breitbart et al., 2004; Breitbart and

Rohwer, 2005; Hurwitz and Sullivan, 2013; Brum et al., 2015), it has its limitations. Viromics typically requires a large sample volume, which may limit the number of samples that can be collected and processed in a timely manner. Moreover, filtration of water samples to physically separate the viral fraction from other microorganisms may lead to the loss of a fraction of the viral community, in particular of giant viruses, while some small bacteria might pass through the filter (Martinez Martinez et al., 2014).

The aim of this study was to assess the composition of the viral communities along the deep-water masses of the thermohaline circulation using a targeted viromics approach. Seawater samples from the core of the NADW (~ 2500 m) and the Pacific AABW (~ 4000 m) were used to discriminate and sort virus-like particles (VLPs) by flow cytometry. The sorted particles were subsequently used to produce DNA libraries for next-generation sequencing (Martinez Martinez et al., 2014). Our analyses provide insight into the composition of viral communities through the circulation path of two major deep-water masses flowing from the Atlantic Ocean through the Southern Ocean to the Pacific. Our data suggest that the thermohaline circulation together with the flux of sinking particles from the epipelagic layer shape the viral community of the deep ocean.

MATERIALS AND METHODS

Study Area and Sampling

North Atlantic Deep Water (~ 2500 m) samples were collected at nine stations during the GEOTRACES cruises 1–3 (2010–2011) in the Atlantic Ocean, on board of the R/V *Pelagia* and R/V *James Cook*. Antarctic Bottom Water (~ 4000 m) samples were collected at seven stations during the KH13-7 and KH14-3 (2013–2014) cruises in the Pacific Ocean on board of the R/V HAKUHO MARU (Figure 1). Sampling was performed using a rosette water sampler equipped with 12 or 24 Niskin bottles with the capacity of either 12 or 25-L, a CTD system (conductivity-temperature-depth; SBE43 Seabird, Bellevue, WA, United States), and a dissolved oxygen sensor.

Two sets of 2-mL water samples from each depth and station were fixed with 0.5% (Atlantic samples) and 1% glutaraldehyde (Pacific samples) held at 4°C for 30 min and flash-frozen in liquid nitrogen thereafter. The samples were stored at -80°C until further processing.

Determination of Microbial Abundance and Sorting of Viruses

Flow cytometry following nucleic acid staining (SYBR Green I) was used to enumerate both, viruses and microbes as described elsewhere (Del Giorgio et al., 1996; Marie et al., 1999; Brussaard, 2004). According to their respective signature in the cytogram of green fluorescence vs. side scatter, two different viral clusters (high and medium fluorescence VLPs) were discriminated (Supplementary Figure S1).

A second set of glutaraldehyde-fixed samples was used for sorting VLPs. After thawing on ice, the samples were rinsed three times with 1xTE Low-EDTA buffer (10 mM Tris-HCl,

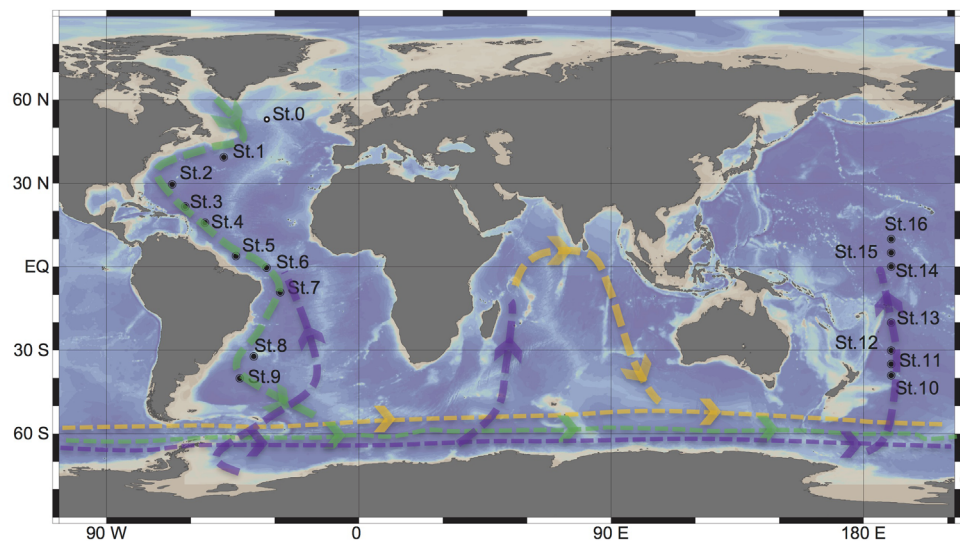


FIGURE 1 | Sampling stations for targeted viromic analysis occupied during the GEOTRACES research expeditions 1, 2 and 3 (2010–2011) in the Atlantic (St.1 to 9) and during KH13-7 and the KH14-3 cruises (2013–2014) in the Pacific (St.10 to 16). The lines represent the simplified scheme of the deep thermohaline circulation, modified from Talley (2013): Green = North Atlantic Deep Water (NADW), purple = Antarctic Bottom Water (AABW), Yellow = Indian Deep Water (IDW). St.0 sampled during the MEDEA2 cruise and used for standard viromic analysis is also shown.

0.1 mM EDTA, pH 8.0) using Amicon 3 kDa centrifuge filters (Millipore) to remove excess glutaraldehyde, which may prevent or decrease the efficiency of downstream DNA amplification (Das et al., 2014). Sorting of viruses was conducted at the JJ MacIsaac Facility for Aquatic Cytometry of the Bigelow Laboratory for Ocean Sciences, with an Influx flow cytometer (BD Biosciences) following an established protocol (Martinez Martinez et al., 2014). Briefly, SYBR Green I stained viruses were discriminated by triggering on side scatter and sorted according to their signature in green fluorescence and side scatter signals. The gain on the 531 nm (green) emission photomultiplier (PMT) was 42.21 and the threshold was set at 0.18. A blank to determine the background noise level was run prior to the samples, and the sorting gates were determined to ensure particle sorting above the background signal region (**Supplementary Figure S1**). The “single 1 drop” mode was used to avoid non-target particles in the target particle droplet and the adjacent droplets. Approximately 6,000 to 10,000 virus particles from each cluster were sorted into 1.5 mL LoBind, sterile microcentrifuge tubes (Eppendorf) for viromics analysis.

Whole Genome Amplification of Sorted Viral DNA and Next-Generation Sequencing (Sorted Viromes)

Aliquots containing between ~1,000 and ~5,000 sorted virus particles (detailed information in **Supplementary Table S1**) from each viral cluster (high and medium nucleic acid fluorescence virus particles) were drawn from each microcentrifuge tube and subjected to whole genome amplification using Illustra Single Cell GenomiPhi DNA Amplification kit (GE Life Sciences) following the manufacturer’s protocol. Subsequently, the amplified genetic material was purified with a PCR purification kit (Qiagen),

quality checked by 2% agarose gel electrophoresis and quantified using a Qubit fluorometer (ThermoFisher). The DNA libraries were constructed using KAPA hyper DNA Library Preparation Kit (with thermocycling consisting of 12 cycles) following the manufacturer’s recommendations and sequenced using Illumina MiSeq high throughput sequencing (2 × 250 bp paired-end platform).

Standard Viromes

Seawater samples were collected at Station 0 in the North Atlantic at 25 m (220 L), the oxygen minimum zone (OMZ) [460 m], 260 L) and 2,000 m (460 L) during the MEDEA-2 cruise in July 2012 (**Figure 1**). The seawater samples were serially pre-filtered through 1.0 and 0.2 μm polycarbonate filters (Millipore) to remove most microorganisms and to enrich the viral fraction. Subsequently, the <0.2 μm filtrate, containing the viral particles, was concentrated by tangential flow filtration using a 100 kDa cutoff cartridge (Millipore) to a final volume of ~150 mL. The concentrates were stored at –80°C until further processing. Nucleic acids from the concentrates were extracted and purified using the phenol/chloroform protocol. The quality and quantity of the extracted DNA was checked by a Nanodrop spectrophotometer (ThermoFisher Scientific). Standard virome DNA libraries were prepared with NEBNext DNA Library Prep Master Mix with thermocycling consisting of 12 cycles, and sequenced at Eurofins Genomics using an Illumina HiSeq high throughput sequencer (2 × 100 bp paired-end platform).

Bioinformatic Analyses

The obtained reads were quality trimmed and controlled using trimmomatic (Bolger et al., 2014) with Phred33. Reads shorter than 36 bp were removed from the analysis. The reads

from the sorted viromes were first digitally normalized using *bbmap*¹, and subsequently taxonomically classified with the Kaiju classifier (Menzel et al., 2016) using the non-redundant NCBI database that included viruses, prokaryotes, fungi and microbial eukaryotes.

Assembly of the filtered reads was performed using *metaSpades* (Nurk et al., 2017) with 21, 33, 55 k-mers size. The obtained contigs were mined for viral signal by *VirSorter* (Roux et al., 2015) implemented in *iVirus*² (Bolduc et al., 2017) against their virome database (Roux et al., 2015) and Viral Single Amplified Genomes (vSAGs) obtained from Martinez-Hernandez et al. (2017). Only phage contigs (from category 1 to 3) larger than 1,500 bp were retained for further analysis. The viral contigs are publicly available in the iMicrobe platform³. To create a read count per contig table, the reads were mapped back to the obtained viral contigs using *Bowtie 2* (local alignment, sensitive setting) (Langmead and Salzberg, 2012) implemented in *iVirus* (Bolduc et al., 2017). *Anvio* v.5.2 (Eren et al., 2015) was used to generate coverage profiles of the mapped reads into viral contigs using only the second and third quartile coverage value from each contig. Open reading frames (ORFs) of the viral contigs were identified using *Prodigal* (Hyatt et al., 2010). The ORFs taxonomic affiliation was determined with *BLASTx* (Altschul et al., 1997) with an *e*-value cut-off of $<10^{-5}$ using *Uniref 100* viral database. Contigs affiliated with the Taterapox virus and *Propionibacterium* phages were discarded from the analysis as potential laboratory contaminants (Woyke et al., 2011; Martinez Martinez et al., 2014). Additionally, the obtained contigs were classified by *Genome-BLAST Distance Phylogeny* (GBDP) (Meier-Kolthoff et al., 2013) using the settings recommended for prokaryotic viruses (Meier-Kolthoff and Goker, 2017). The evolution trees obtained from the intergenomic distances were calculated on *FASTME* (Lefort et al., 2015) with branch support inferred from 100 pseudo-bootstraps. The obtained trees were rooted at the midpoint (Farris, 1972) and visualized with *FigTree* (Rambaut, 2006). Taxon classification was estimated with the *OPTSIL* program (Goker et al., 2009) with the recommended clustering thresholds (Meier-Kolthoff and Goker, 2017) and an *F* value of 0.5 (Meier-Kolthoff et al., 2014).

The sequence data generated are publicly available in the DDBJ database under the accession number DRA008218.

Viral Recruitment Analysis

The sorted and standard viromes were compared to five viral genomes and four environmental samples (three vSAGs and one assembled genome) isolated from marine ecosystems. From each virome, 1.5 million randomly subsampled reads were recruited against reference sequences using *tBLASTx* with an *e*-value cut-off of $<10^{-5}$. Both the average and percentage coverage values were calculated from the nucleotide alignment positions of the reads (obtained from the *tBLASTx* results) against the reference genome.

Protein Clustering Analysis

The viral contigs were quality controlled using *QUAST*, allowing one ambiguity and the threshold for misassemblies set to 1000 bp (Gurevich et al., 2013). The ORFs were predicted and annotated using *PROKKA* (Seemann, 2014) optimized in house for virus detection. Briefly, *PROKKA* was supplemented with a *prodigal* (Hyatt et al., 2010) in house-trained HMM model of the non-redundant Reference Viral Database (RVDB version April 2018, Goodacre et al., 2018) containing viral, virus-like and virus-related sequences. The predicted ORFs were clustered to existing proteins obtained from the Reference Viral (Goodacre et al., 2018) and Tara Ocean Database (Brum et al., 2015) using *CD-HIT-2D* (Li and Godzik, 2006) with a sequence cutoff of 60% and a coverage cutoff of 80%. To create a table with the read counts per protein clusters, the reads were mapped back to the obtained ORFs using *Bowtie 2* (Langmead and Salzberg, 2012).

Statistical Analysis

Analysis of variance (ANOVA test) was performed to test possible variations between each water mass. The richness of the viral contigs and viral protein clusters were calculated from the coverage table using “*Vegan*” package (Oksanen et al., 2018) implemented in R software v.3.4. Jaccard Dissimilarity index was calculated from the subsampled (200,000 reads per sample) coverage table (obtained from the viral contigs and protein clusters) using the “*Vegan*” package.

The presence of specific protein clusters in the different samples was analyzed with the “*Venn*” package in the R software using presence/absence data.

RESULTS

Environmental and Microbial Parameters

The NADW and the Pacific Antarctic Bottom Water (PAABW) water differed in their physico-chemical characteristics. NADW exhibited higher temperature, salinity and dissolved oxygen concentrations ($2.9 \pm 0.5^{\circ}\text{C}$, 34.9 ± 0.1 , $233 \pm 13 \mu\text{mol/kg}$, respectively) than the PAABW ($1.3 \pm 0.13^{\circ}\text{C}$, 34.7 ± 0.0 , $172 \pm 23 \mu\text{mol/kg}$, respectively) (Table 1). Prokaryotic and viral abundance were significantly higher in the NADW ($1.4 \pm 0.9 \times 10^4 \text{ mL}^{-1}$, $5.2 \pm 2.0 \times 10^5 \text{ mL}^{-1}$, respectively) than in the PAABW, with prokaryotic and viral abundance of $0.7 \pm 0.1 \times 10^4 \text{ mL}^{-1}$ and $1.2 \pm 0.2 \times 10^5 \text{ mL}^{-1}$, respectively (*T*-test, $p < 0.05$). The viral to prokaryotic ratio (VPR) was also significantly higher in the NADW (45 ± 18) than in the PAABW (16 ± 2) (*T*-test, $p < 0.05$).

Taxonomic Composition of the Sorted Viromes

Unclassified reads accounted for $52 \pm 22\%$ of the total sequences, followed by Bacteria ($40 \pm 23\%$), Eukarya ($4 \pm 2\%$), viruses ($3 \pm 2\%$) and Archaea ($0.6 \pm 0.2\%$). Consequently, excluding the reads with no or unknown homology, viruses accounted for $\sim 7\%$ of the DNA library. The North Atlantic and North Pacific libraries comprised the lower proportion of viral reads

¹<https://jgi.doe.gov/>

²<https://www.cyverse.org/>

³<https://www.imicrobe.us/#/projects/334>

TABLE 1 | Physico-chemical and biological characteristics of the North Atlantic Deep Water (NADW) and Pacific Antarctic Bottom Water (PAABW) sampled in the Atlantic and Pacific Ocean, respectively.

Samples	N	Water mass	Depth (m)	Environmental Parameters				Microbial Counts			
				Temperature (°C)	Salinity	Oxygen (μmol/kg)	Prokaryotic Abundance ($\times 10^4$ mL ⁻¹)	Viral Abundance ($\times 10^5$ mL ⁻¹)	VPR	Average $\pm$ SD	Average $\pm$ SD
				Average $\pm$ SD	Average $\pm$ SD	Average $\pm$ SD	Average $\pm$ SD	Average $\pm$ SD	Average $\pm$ SD		
Atlantic	4	NADW	2500	3.1 $\pm$ 0.2	34.9 $\pm$ 0.0	234 $\pm$ 3.4	0.8 $\pm$ 0.2	4.8 $\pm$ 1.1	61 $\pm$ 5		
Equator	2	NADW	~2500	2.9 $\pm$ 0.1	34.9 $\pm$ 0.0	240 $\pm$ 13.2	1.5 $\pm$ 0.8	3.6 $\pm$ 1.3	29 $\pm$ 23		
South Atlantic	3	NADW	2500	2.5 $\pm$ 0.7	34.9 $\pm$ 0.1	227 $\pm$ 23	2.2 $\pm$ 1.2	7.0 $\pm$ 2.3	34 $\pm$ 7		
South Pacific	4	AABW	~4000	1.2 $\pm$ 0.1	34.7 $\pm$ 0.0	191 $\pm$ 3.6	0.8 $\pm$ 0.1	1.3 $\pm$ 0.2	17 $\pm$ 2		
Equator	2	AABW	~4000	1.4 $\pm$ 0.0	34.7 $\pm$ 0.0	146 $\pm$ 3.0	0.8 $\pm$ 0.2	1.2 $\pm$ 0.1	17 $\pm$ 1		
North Pacific*	1	AABW	~4000	1.4	34.7	150	0.5	0.8	15		

*A single data point was collected at this station. Therefore, average and standard deviation are not reported. Average $\pm$ standard deviation (SD) and number of samples (N) are indicated.

with 2% and 1% of their corresponding libraries, respectively. In contrast, viral sequences contributed 12% and 10% to the DNA libraries of the South Atlantic and Equatorial Pacific, respectively (**Supplementary Figure S2**).

Characteristics of the Sorted Viromes

Thirty-two sorted viromes were produced from seawater samples collected at 16 stations (two virus particle clusters sorted from each station) along the global deep ocean conveyor belt following the NADW and the PAABW (**Figure 1**). The viromes were grouped in six geographic regions according to water mass (NADW and PAABW) and location where the samples were collected (northern hemisphere, equator and southern hemisphere), resulting in six virome datasets (**Supplementary Table S1**). The whole genome amplification of sorted viruses yielded a total of 43 million quality-trimmed reads, assembled into 575 viral contigs (205 contigs from NADW and 370 contigs from PAABW) as predicted with VirSorter, with a minimum length of ~1,500 bp and a maximum length of ~18,000 bp (**Table 2**). Fifteen percent of the total quality-trimmed reads were mapped to the obtained viral contigs. These represented 12% of the reads from NADW and 18% of the PAABW reads (**Table 2**).

Relative Abundance and Diversity of Viral Communities

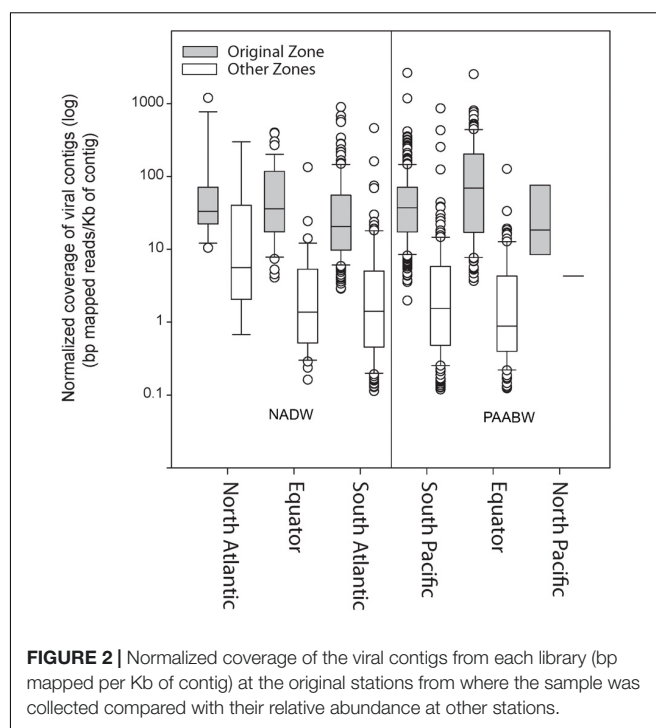
Read recruitment from any of the six combined viromes to itself was higher (on average 92 bp mapped per kilobase of contigs) than the recruitment to viromes from any of the other water mass and region (11 bp mapped per kilobase of contigs) (**Figure 2**).

The number of viral protein clusters shared between the NADW and PAABW in different oceanic regions is depicted in **Figure 3**. The NADW and the PAABW shared 661 protein clusters (18% of total protein clusters) (**Figure 3A**). PAABW harbored a larger number of unique protein clusters (1870 protein clusters, 52% of the total protein clusters) than the NADW (1078 protein clusters, 30%) (**Figure 3A**). NADW viromes from the North Atlantic, Equator and South Atlantic shared only 14 protein clusters (0.8% of NADW protein clusters) (**Figure 3B**). The South Atlantic NADW harbored the highest number of unique protein clusters, followed by the NADW at the Equator and the North Atlantic (**Figure 3B**). PAABW viromes shared only 5 protein clusters (0.2%) between the three regions of the Pacific Ocean (**Figure 3B**). Similarly, to the NADW, PAABW harbored the largest number of unique protein clusters in the South Pacific followed by the Equator and North Pacific (**Figure 3B**).

Jaccard Dissimilarity analysis of the viral communities and protein clusters along the NADW and PAABW further confirmed that the viral communities differed between different regions and water masses (**Figures 4A,B**). Adjacent regions were generally more similar to each other than to more distant regions, i.e., a decreasing similarity can be observed between North Atlantic, the Equator and South Atlantic (**Figures 4A,B**) and from South Pacific northwards. The viral communities from the southern hemisphere clustered with the equatorial Pacific, while the viral

TABLE 2 | Summary of the total number of contigs (N) (obtained with Virsorter, minimum length 1,500 bp), their average (avg_len), minimum (min_len) and maximum (max_len) length (bp), and %GC content of the contigs. Total number of trimmed reads (num_seqs) and their length (sum_len). Total number of mapped reads into viral contigs (num_seqs), their length (sum_len) and the % of mapped reads to the viral contigs are also indicated.

Samples	Contigs					Reads				
	Viral contigs mined with Virsorter					Trimmed Reads				
	N	min_len (bp)	avg_len (bp)	max_len (bp)	GC%	num_seqs	sum_len (bp)	num_seqs	sum_len (bp)	% mapped reads
Atlantic	North Atlantic	11	1,507	4,146	38	13,944,333	3,390,243,489	1,466,305	270,703,351	11
	A_Equator	48	1,673	2,793	37	5,076,934	1,134,402,129	615,029	123,249,466	12
Pacific	South Atlantic	146	1,509	3,517	36	7,319,179	1,832,146,844	983,162	199,166,371	13
	South Pacific	244	1,503	3,910	36	9,229,727	2,516,721,586	2,032,162	505,268,561	22
	P_Equator	122	1,541	3,496	36	6,073,330	1,832,829,353	1,318,215	340,742,928	22
	North Pacific	4	3,636	6,216	47	1,959,466	423,056,799	216,239	45,300,028	11



communities in the northern hemisphere clustered with the equatorial Atlantic (**Supplementary Figures S3A,B**).

Viral Communities Distribution Along the Deep-Water Thermohaline Circulation

Pronounced regional differences of the viral communities were detected along the deep-water circulation with most of the reads from any given virome mapping to the contigs from the same (“original”) virome (**Figure 5**). Some viral contigs were also ubiquitously distributed across both water masses (**Figure 5** and **Supplementary Table S2**). Noticeably, a large fraction of the contigs obtained from the deep sorted viromes mapped to reads obtained from the surface standard virome sampled at St.0 (**Figure 5**). Most of the viral contigs of the NADW and PAABW were affiliated with uncultured viruses (196 contigs, 34% of the total contigs), an uncultured Mediterranean phage (123 contigs, 21%), followed by *Pelagibacter* phage (30 contigs, 5%), *Synechococcus* phage (28 contigs, 5%), uncultured marine viruses (17 contigs, 3%) and *Prochlorococcus* phage (13 contigs, 2%). Other viral groups affiliated with phages of *Bacillus*, *Cyanobacteria*, *Staphylococcus*, *Caulobacter* and *Gordonia* accounted for ~5% of the total contigs, unassigned and other classified viruses accounted for 16 and 6% of the total viral contigs, respectively. Overall, only 22% of the viral contigs could be affiliated to cultured viruses, mainly bacteriophages. The most abundant cultured viruses found in our database were related to viruses infecting widespread hosts such *Pelagibacter*, *Synechococcus*, and *Prochlorococcus*. The VICTOR-based taxonomic classification of the viral contigs (**Figure 5**, **Supplementary Figure S4**, and **Supplementary Table S3**) exhibited similar trends, with the majority of the virus sequences closely related to uncultured environmental viruses

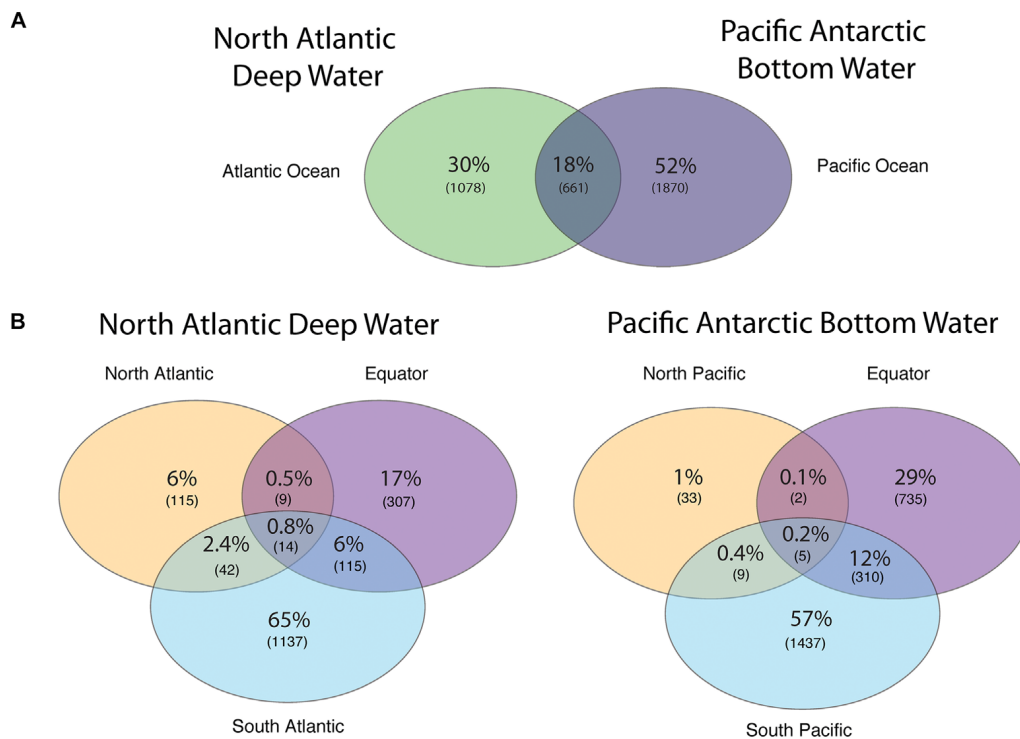


FIGURE 3 | Venn diagram showing the shared and unique viral PCs obtained from subsampled reads (200,000 reads per virome) recruited into viral ORFs **(A)** between the NADW and the PAABW, and **(B)** between different regions through the thermohaline circulation of NADW and of PAABW.

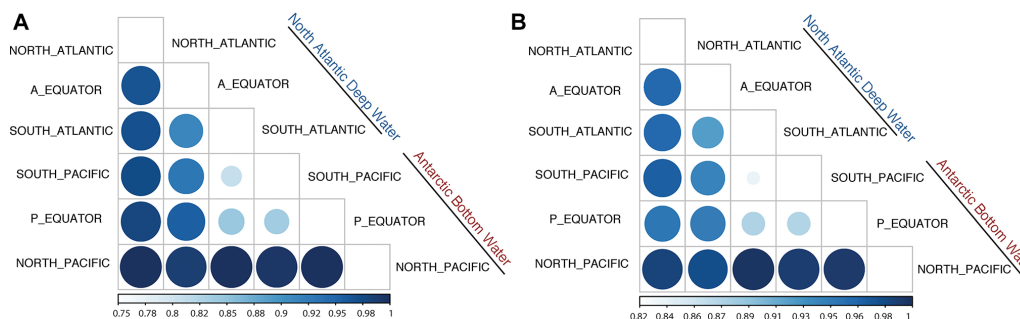


FIGURE 4 | Jaccard Dissimilarity analysis of the subsampled (200,000 reads) sorted viromes **(A)** and PCs **(B)** collected in NADW and PAABW along the deep-water circulation.

(43%) and vSAGs (52%), followed by widespread cultured marine phages (3.5%) (**Figure 5**, Reference 1 Victor-Based Taxonomy).

A significant number of reads was tentatively affiliated with *Prochlorococcus* phages with, on average, 30 bp mapped per kilobase of contigs followed by *Synechococcus* and *Pelagibacter* phages, with 27 and 21 bp mapped per kilobase of contigs, respectively (**Figure 5** and **Supplementary Figure S5**). Sequences tentatively affiliated with an uncultured Mediterranean Sea phage also accounted for a significant proportion of mapped reads (69 bp mapped per kilobase of contigs, on average). Nevertheless, most of the sampled viral groups lacked cultured representatives or viral reference genomes (**Figure 5** and **Supplementary Figure S5**).

Since most of the sorted virome sequences were related to specific viral genomes or environmental virus clones according to the results from the taxonomic analyses, these genomes and clones were selected for the fragment recruitment analyses of our viral reads. The recruitment analyses indicated a low coverage and percentage of similarity between the viral reads obtained in the North Atlantic, Atlantic Equator and North Pacific regions and the reference genomes of the most abundant widespread marine viruses. The larger genome coverage to cultured representatives was found in the South Atlantic, South Pacific and Pacific Equator sorted viromes and in the standard viromes (surface, OMZ and 2000 m) (**Figure 6** and **Supplementary Table S4**). Most of the high coverage

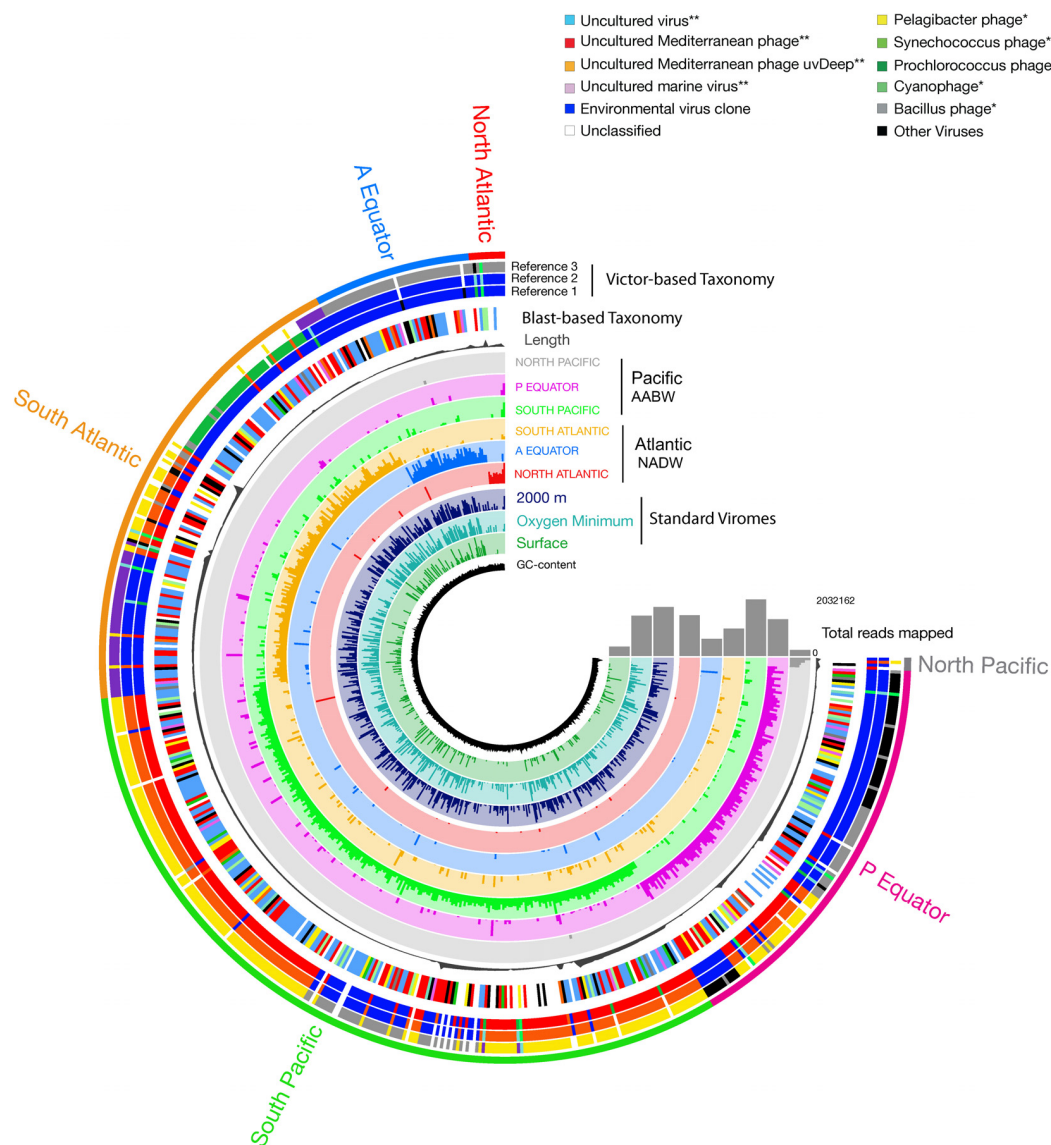


FIGURE 5 | Coverage of viral contigs in NADW and PAABW samples. Coverage in the three standard virome samples from the North Atlantic surface, oxygen minimum and 2000 m is indicated in 2–4 inner rings, respectively. Length of the contigs (gray bar, in bp). The height of the colored bars within a ring represents the mean coverage (second and third quartile) of a contig in a given sample. The samples were ordered by geographic location following the deep-water conveyor belt circulation. The first outer ring indicates the taxonomic classification of the viral contigs based on BLASTx at class (**), and at the genus (*) level. The next 2 to 4 outer rings indicate the taxonomic classification based on VICTOR, showing the three most relevant identified references that clustered with the viral contigs.

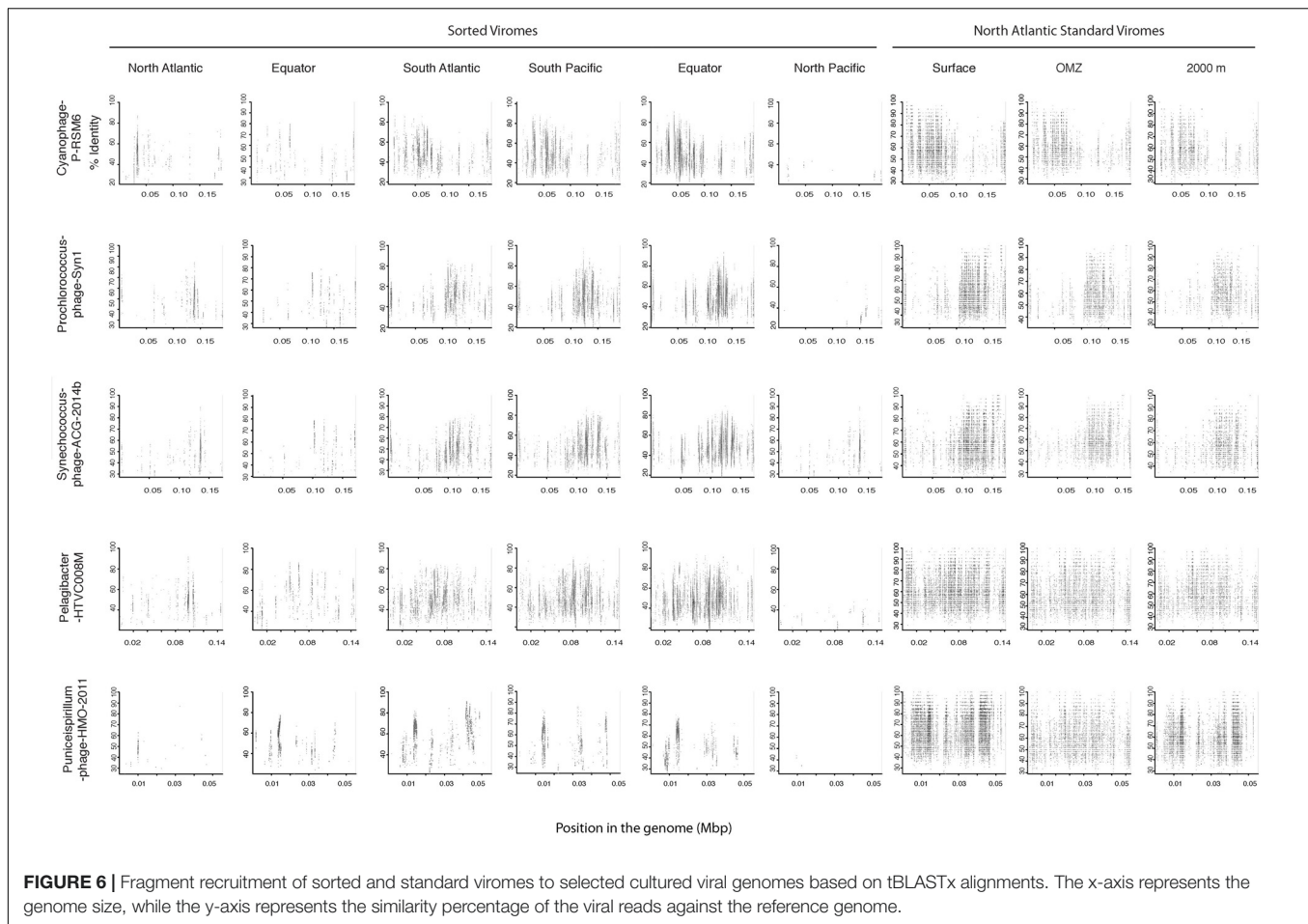
recruitments from the sorted viromes exhibited 40–80% amino acid identity, while similarity greater than 90% was only found within the standard viromes reads recruited against Pelagibacter phage HTVC008M and Puniceispirillum phage HMO-2011 (Figure 6 and Supplementary Table S4). Besides cultured representatives, viral SAGs from surface waters and uncultured marine viruses also showed a relatively high coverage to viral reads recovered from the sorted and standard viromes. The recruitment plot showed that vSAG-37-F6 and vSAG-37-G23 had the highest genome coverage within the sorted and standard viromes (Figure 7 and Supplementary Table S4), while uvDeep-CGR2-KM22-C255 showed the highest coverage and amino

acid similarity in the 2000 m standard virome (Figure 7 and Supplementary Table S4).

DISCUSSION

Biotic and Environmental Variables

The physico-chemical characteristics of the two water masses studied here remained constant, however, the microbial and viral abundance showed some variability, suggesting that particle flux from the photic layer or other aspects, such as *in situ* production, may locally influence the deep microbial and viral



communities (Nagata et al., 2000). Additionally, the NADW, as younger water mass than the PAABW, harbored higher microbial and viral abundances than the PAABW. Thus, with aging of the water masses, microbial and viral abundance are decreasing along the meridional overturning circulation following the general concentration pattern of dissolved organic carbon (DOC) (Hansell, 2013).

Yang et al. (2014) showed that viral abundance significantly decreased from the relatively younger water mass (PAABW) to the older water mass (Pacific Deep Water), coinciding with a 0.4% of the total DOC sink in the Pacific Ocean. The pattern of the virus-to-prokaryote ratio along NADW and PAABW, together with the long turnover time of microorganisms and viruses support a DOC-like distribution of the viruses in the deep water masses. Thus, the transition from younger to older water mass might also have an effect on the viral-host interactions and virus diversity.

Viral Diversity and Distribution Across Two Deep-Water Masses

Most of the identified viral groups were widely distributed along different regions of the thermohaline circulation, although their relative abundance was higher in the region where they were

isolated (Figure 5). This finding suggests a relatively high local variability within a limited global diversity, in agreement with results from other marine viral metagenomes (Hurwitz and Sullivan, 2013; Brum et al., 2015) and a natural *Synechococcus* population (Deng et al., 2014). The low temperature of deep-sea water masses and the consequently longer turnover time of bacteria and slower viral decay in this realm may favor the dispersion of locally produced virus particles across the deep overturning circulation and support the viral seed-bank hypothesis (Breitbart and Rohwer, 2005). Under these environmental conditions, viruses may behave conservatively and thus, might reflect to some extent the different formation sites of the respective deep-water masses. Therefore, the viral composition in the PAABW might reflect the mixture of NADW, Indian Ocean Deep Water and circumpolar deep water mixing and sinking at the Antarctic shelf. In contrast, the viral composition in the NADW may initially reflect the mixture of subtropical surface waters and Arctic shelf waters (Talley, 2013) and through its way toward Antarctic waters, it probably incorporates virus from other sources through mixing with other water masses or particles sedimenting into these water masses. The overlap between surface and deep water viral communities might indicate a transport of viruses associated with sinking particles (Figure 5 and Supplementary

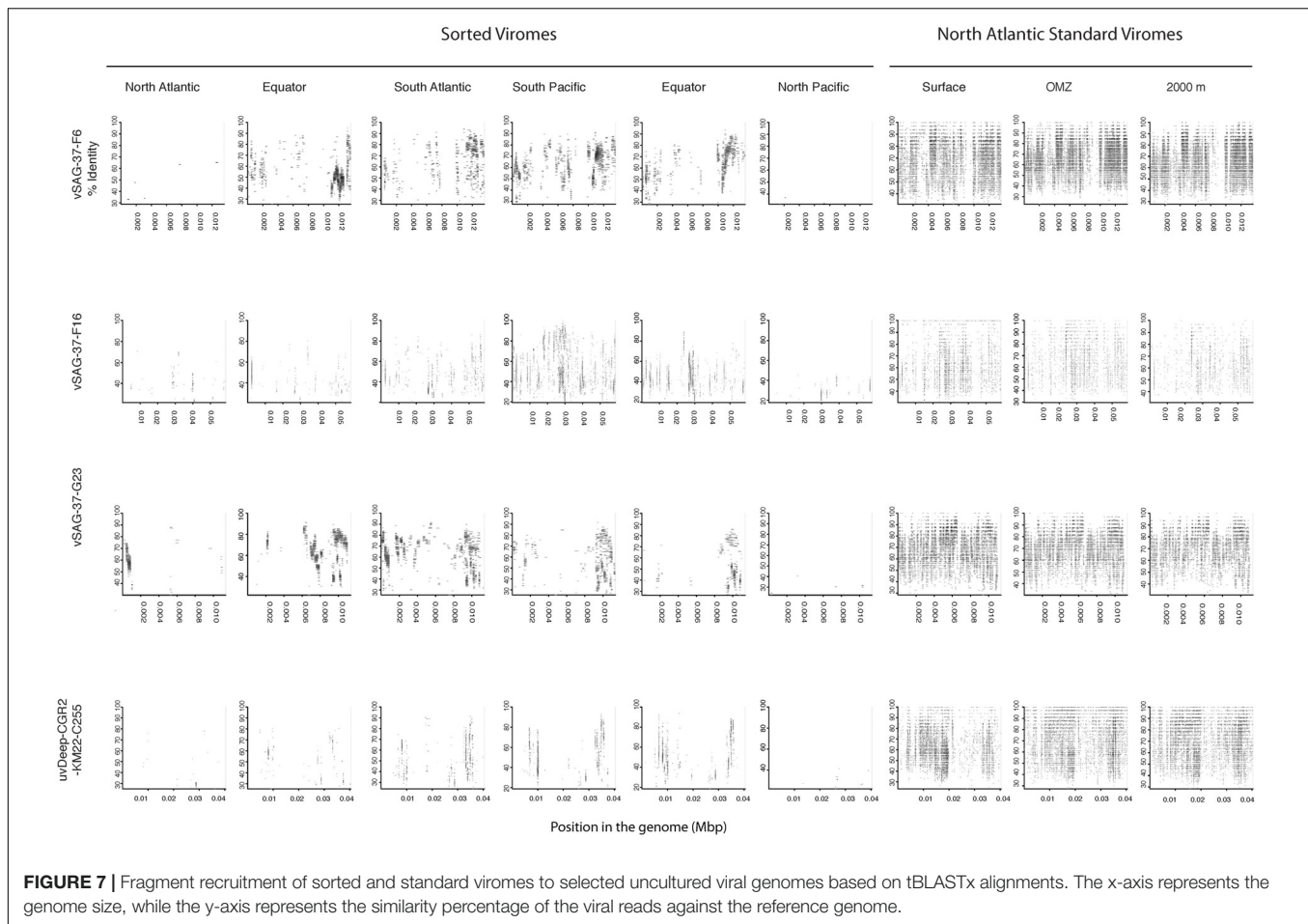


Figure S6) across otherwise stratified waters, suggesting that the deep viral community might partially comprise surface representatives.

The widespread occurrence of some viral groups along the deep thermohaline circulation might be the result of the ubiquitous distribution of key prokaryotic phylotypes found in the deep ocean (Martin-Cuadrado et al., 2007; Sintes et al., 2013).

Archaea represent a large fraction of the microbial community in meso- and bathypelagic waters, reaching up to 30% of the prokaryotic community (Herndl et al., 2005). Recent studies have indicated that Archaea are mainly controlled by viruses in the deep seafloor, suggesting that archaeal phages play an important role in the deep-sea ecosystem (Danovaro et al., 2015, 2016; Nigro et al., 2017). Other bacterial groups such as SAR324, SAR406 also dominated the bathypelagic microbial communities (Lopez-Garcia et al., 2001; Nunoura et al., 2015). Therefore, those viruses infecting the dominant groups should be present. However, viruses infecting the main bathypelagic communities such as Archaea, SAR324 and SAR406 were not detected in NADW and PAABW.

Only a small proportion of the obtained reads were mapped to the obtained viral contigs. The scarcity of identified representatives of deep water viruses also suggests that one of the main limitations in assessing the deep ocean viral community

structure is the low number of both microbial and viral genomes available from this environment (Mizuno et al., 2016).

Sinking Particles Influence Microbial and Viral Community Composition in Deep Waters

Viral contigs associated with phages of *Synechococcus* and *Prochlorococcus* comprised a considerable fraction of viral contigs in the NADW and PAABW. Additionally, the similarity observed between deep-ocean viral reads and widespread bacteriophages infecting photosynthetic hosts (**Figure 6**), and the high coverage and amino acid similarity between surface and deep viral contigs (**Supplementary Figure S6**) support the idea of an active transport of microbes and viruses from the epipelagic realm through sinking particles as previously reported (Hurwitz et al., 2014, 2015) based on the presence of photosystem II reaction center genes (*psbA*) in bathypelagic viruses. In agreement with our findings, Gregory et al. (2019) suggested that the deep-sea viral populations may be more dependent on migration from sinking habitats rather than on local production.

The high virus-to-prokaryote ratio found in deep waters (Yang et al., 2014; Lara et al., 2017) also points to the transfer of viruses originated from the surface ocean to the dark ocean through

sinking particles. It is unlikely that phages of cyanobacteria will find a host in the deep ocean. Hence, an ecological function of these phages is difficult to envision in that environment. However, it cannot be excluded that wide-spread viral groups and/or broad host range viruses originating from the ocean surface contribute to the mortality of deep sea microbes and, hence, affecting the carbon and nutrient cycling in the meridional overturning circulation.

It is likely that *in situ* viral production through lytic infection of abundant and active heterotrophic microbes, supported by organic sinking particles, also contribute to the high virus-to-prokaryote ratios we measured. Nevertheless, our sequencing data did not show any clear evidence concerning the life strategies of deep-sea viruses.

In aged water masses such as the North PAABW, where the available DOC is old and refractory (Hansell, 2013), sinking particles may provide the only source of organic matter available for the activity of the deep-sea microbial community (Hansell and Ducklow, 2003) and, consequently, shaping the viral communities.

CONCLUSION

Sorted viromes were used to assess the changes in viral community across two major water masses NADW and PAABW of the deep thermohaline circulation. Significant differences in viral abundance and composition were found between the NADW and PAABW. The proportion of shared viral protein clusters between the two water masses together with the presence of widespread cultured and uncultured viral groups suggest that the viral distribution along the deep overturning circulation may be at least partially explained by passive dispersion. This would confirm the seed-bank hypothesis. Nevertheless, the presence of viral groups with phototrophic hosts in these bathypelagic waters also suggests that sinking particles may play an important role in shaping the deep-sea microbial community and consequently, the viral community. Therefore, sinking particles influence the virus-host interactions and the biogeochemical cycles of the deep ocean. The lack of deep-sea viral and microbial representatives remains one of the main bottlenecks in this research. It prevents the identification of the key viruses and the assessment of viral-host interactions in the dark ocean.

DATA AVAILABILITY

The datasets generated for this study can be found in the DDBJ database under the accession number DRA008218.

AUTHOR CONTRIBUTIONS

DC, JM, and TY contributed to the concept and design of the study and wrote the manuscript. MC and YT performed the part of the bioinformatic analyses. DC analyzed the data and wrote the first draft of the manuscript. GH,

ES, and TN provided feedback and contributed to the manuscript revision. DC, GH, and TN provided funding for the analyses.

FUNDING

Laboratory work was supported by the JSPS KAKENHI (Grant Numbers JP15K12189 and JP16F16085) and grants for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Science, Sports and Technology (MEXT) of Japan (Grant Numbers 16H06429, 16K21723, and 16H06437). DC was supported by an Overseas Researcher Grant under Postdoctoral Fellowship of Japan Society for Promotion of Science (P16085). ES was supported by the Austrian Science Fund (FWF) project P27696-B22. JM was supported by the Gordon and Betty Moore Foundation grant GBMF5334. MC support for this project was provided through Bigelow Laboratory 442 for Ocean Sciences internal funding to JM, and through NSF-OPP 1644155 awarded to JM. The MEDEA-2 cruise was funded by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement No. 268595 (MEDEA project) to GH.

ACKNOWLEDGMENTS

We thank Miho Hirai for the support during the laboratory experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.01801/full#supplementary-material>

FIGURE S1 | Representative cytograms of green fluorescence (SYBR Green I) versus side scatter and green fluorescence versus forward scatter showing the actual sorting gates (P2 and P3) employed for the two viral populations targeted in this study. Gate P5 in green fluorescence versus forward scatter was used to discriminate background within the sorting gates.

FIGURE S2 | Classification of the reads obtained from the sorted viromes based on Kaiju classification with the non-redundant NCBI database.

FIGURE S3 | Clustering of the sorted viromes (A) and PCs (B) based on the Jaccard Dissimilarity matrix.

FIGURE S4 | Phylogenomic trees (of the sorted viromes) obtained using the Genome-BLAST Distance Phylogeny (GBDP) method, inferred using the formulas D0. The numbers above branches are the pseudo-bootstrap support values from 100 replications. The branch lengths of the resulting VICTOR trees are scaled in terms of the respective distance formula used.

FIGURE S5 | Taxonomic identification and relative abundance (bp mapped per Kb of contig) of the most abundant viral contigs sampled in NADW and PAABW.

FIGURE S6 | Fragment recruitment of the surface standard virome to six selected deep viral contigs (obtained in the NADW and PAABW) based on tBLASTx alignments. The x-axis represents the genome size, while the y-axis represents the similarity percentage of the viral reads against the reference genome.

TABLE S1 | Physico-chemical, microbial and viral parameters sampled at different stations in the Atlantic and Pacific Ocean. All the parameters are indicated for each individual sample. The number of viruses sorted and amplified with the whole genome amplification kit (WGA) and the obtained trimmed reads are also stated. PA, prokaryotic abundance, VA, Viral abundance, VPR, Viral to prokaryotic ratio. Grouping of samples according to oceanic region: NA, North Atlantic, EA, Equatorial Atlantic, SA, South Atlantic, SP, South Pacific, EP, Equatorial Pacific, NP, North Pacific.

TABLE S2 | Percentage of recruited reads against the viral contigs sampled in the NADW and AABW along the deep ocean conveyor belt.

TABLE S3 | Clustering of the viral contigs against selected viral references based on the OPTSIL algorithm.

TABLE S4 | Fragment recruitment results of selected viral genomes and uncultured marine phages used as references against reads obtained from the sorted and standard viromes.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., et al. (2006). The marine viromes of four oceanic regions. *PLoS Biol.* 4:e368. doi: 10.1371/journal.pbio.0040368
- Bercovici, S. K., and Hansell, D. A. (2016). Dissolved organic carbon in the deep southern ocean: local versus distant controls. *Glob. Biogeochem. Cycles* 30, 350–360. doi: 10.1002/2015gb005252
- Bolduc, B., Youens-Clark, K., Roux, S., Hurwitz, B. L., and Sullivan, M. B. (2017). iVirus: facilitating new insights in viral ecology with software and community data sets imbedded in a cyberinfrastructure. *ISME J.* 11, 7–14. doi: 10.1038/ismej.2016.89
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bongiorni, L., Magagnoli, M., Armeni, M., Noble, R., and Danovaro, R. (2005). Viral production, decay rates, and life strategies along a trophic gradient in the North Adriatic Sea. *Appl. Environ. Microbiol.* 71, 6644–6650. doi: 10.1128/aem.71.11.6644-6650.2005
- Breitbart, M., Miyake, J. H., and Rohwer, F. (2004). Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiol. Lett.* 236, 249–256. doi: 10.1016/j.femsle.2004.05.042
- Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.* 13, 278–284. doi: 10.1016/j.tim.2005.04.003
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulier, G., Acinas, S. G., Alberti, A., et al. (2015). Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498. doi: 10.1126/science.1261498
- Brussaard, C. P. D. (2004). Optimization of procedures for counting viruses by flow cytometry. *Appl. Environ. Microbiol.* 70, 1506–1513. doi: 10.1128/aem.70.3.1506-1513.2004
- Cho, B. C., and Azam, F. (1988). Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature* 332:441. doi: 10.1038/332441a0
- Danovaro, R., Corinaldesi, C., Rastelli, E., and Dell'Anno, A. (2015). Towards a better quantitative assessment of the relevance of deep-sea viruses, bacteria and archaea in the functioning of the ocean seafloor. *Aqua. Microbial. Ecol.* 75, 81–90. doi: 10.3354/ame01747
- Danovaro, R., Dell'anno, A., Corinaldesi, C., Rastelli, E., Cavicchioli, R., Krupovic, M., et al. (2016). Virus-mediated archaeal hecatomb in the deep seafloor. *Sci. Adv.* 2:e1600492. doi: 10.1126/sciadv.1600492
- Das, K., Fernando, M. R., Basiaga, S., Wigginton, S. M., and Williams, T. (2014). Effects of a novel cell stabilizing reagent on DNA amplification by PCR as compared to traditional stabilizing reagents. *Acta. Histochemica* 116, 55–60. doi: 10.1016/j.acthis.2013.05.002
- De Corte, D., Sintes, E., Yokokawa, T., and Herndl, G. J. (2011). Changes in viral and bacterial communities during the ice-melting season in the coastal Arctic (Kongsfjorden, Ny-Alesund). *Environ. Microbiol.* 13, 1827–1841. doi: 10.1111/j.1462-2920.2011.02497.x
- De Corte, D., Sintes, E., Yokokawa, T., Lekunberri, I., and Herndl, G. J. (2016). Large-scale distribution of microbial and viral populations in the South Atlantic Ocean. *Environ. Microbiol. Rep.* 8, 305–315. doi: 10.1111/1758-2229.12381
- De Corte, D., Sintes, E., Yokokawa, T., Reinthaler, T., and Herndl, G. J. (2012). Links between viruses and prokaryotes throughout the water column along a North Atlantic latitudinal transect. *ISME J.* 6, 1566–1577. doi: 10.1038/ismej.2011.214
- Del Giorgio, P., Bird, D. F., Prairie, Y. T., and Planas, D. (1996). Flow cytometric determination of bacterial abundance in lake plankton with the green nucleic acid stain SYTO 13. *Limnol. Oceanogr.* 41, 783–789. doi: 10.4319/lo.1996.41.4.0783
- Deng, L., Ignacio-Espinoza, J. C., Gregory, A. C., Poulos, B. T., Weitz, J. S., Hugenholtz, P., et al. (2014). Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* 513, 242–245. doi: 10.1038/nature13459
- Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. doi: 10.7717/peerj.1319
- Farris, J. S. (1972). Estimating phylogenetic trees from distance matrices. *Am. Nat.* 106, 645–668. doi: 10.1086/282802
- Follett, C. L., Repeta, D. J., Rothman, D. H., Xu, L., and Santinelli, C. (2014). Hidden cycle of dissolved organic carbon in the deep ocean. *Proc. Natl. Acad. Sci. U. S. A.* 111, 16706–16711. doi: 10.1073/pnas.1407445111
- Frank, A. H., Garcia, J. A. L., Herndl, G. J., and Reinthaler, T. (2016). Connectivity between surface and deep waters determines prokaryotic diversity in the North Atlantic Deep Water. *Environ. Microbiol.* 18, 2052–2063. doi: 10.1111/1462-2920.13237
- Goker, M., Garcia-Blazquez, G., Voglmayr, H., Telleria, M. T., and Martin, M. P. (2009). Molecular taxonomy of phytopathogenic fungi: a case study in *Peronospora*. *PLoS One* 4:e6319. doi: 10.1371/journal.pone.0006319
- Goodacre, N., Aljanahi, A., Nandakumar, S., Mikailov, M., and Khan, A. S. (2018). A reference viral database (rvdb) to enhance bioinformatics analysis of high-throughput sequencing for novel virus detection. *mSphere* 3, e00069-18. doi: 10.1128/mSphereDirect.00069-18
- Gregory, A. C., Zayed, A. A., Conceicao-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177:e1114. doi: 10.1016/j.cell.2019.03.040
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hansell, D. A. (2013). Recalcitrant dissolved organic carbon fractions. *Annu. Rev. Mar. Sci.* 5, 421–445. doi: 10.1146/annurev-marine-120710-100757
- Hansell, D. A., and Ducklow, H. W. (2003). Bacterioplankton distribution and production in the bathypelagic ocean: directly coupled to particulate organic carbon export? *Limnol. Oceanogr.* 48, 150–156. doi: 10.4319/lo.2003.48.1.0150
- Herndl, G. J., Reinthaler, T., Teira, E., Van Aken, H., Veth, C., Pernthaler, A., et al. (2005). Contribution of archaea to total prokaryotic production in the deep atlantic ocean. *Appl. Environ. Microbiol.* 71, 2303–2309. doi: 10.1128/aem.71.5.2303-2309.2005
- Hurwitz, B. L., Brum, J. R., and Sullivan, M. B. (2015). Depth-stratified functional and taxonomic niche specialization in the 'core' and 'flexible' Pacific Ocean Virome. *ISME J.* 9, 472–484. doi: 10.1038/ismej.2014.143
- Hurwitz, B. L., and Sullivan, M. B. (2013). The pacific ocean virome (pov): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS One* 8:e57355. doi: 10.1371/journal.pone.0057355

- Hurwitz, B. L., Westveld, A. H., Brum, J. R., and Sullivan, M. B. (2014). Modeling ecological drivers in marine viral communities using comparative metagenomics and network analyses. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10714–10719. doi: 10.1073/pnas.1319778111
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Karl, D. M., Knauer, G. A., and Martin, J. H. (1988). Downward flux of particulate organic matter in the ocean: a particle decomposition paradox. *Nature* 332:438. doi: 10.1038/332438a0
- Kukkaro, P., and Bamford, D. H. (2009). Virus-host interactions in environments with a wide range of ionic strengths. *Environ. Microbiol. Rep.* 1, 71–77. doi: 10.1111/j.1758-2229.2008.00007.x
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lara, E., Vaque, D., Sa, E. L., Boras, J. A., Gomes, A., Borrull, E., et al. (2017). Unveiling the role and life strategies of viruses from the surface to the dark ocean. *Sci. Adv.* 3:e1602565. doi: 10.1126/sciadv.1602565
- Lefort, V., Desper, R., and Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* 32, 2798–2800. doi: 10.1093/molbev/msv150
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Lopez-Garcia, P., Lopez-Lopez, A., Moreira, D., and Rodriguez-Valera, F. (2001). Diversity of free-living prokaryotes from a deep-sea site at the antarctic polar front. *FEMS Microbiol. Ecol.* 36, 193–202. doi: 10.1016/s0168-6496(01)00133-7
- Marie, D., Brussaard, C. P. D., Thyrhaug, R., Bratbak, G., and Vaulot, D. (1999). Enumeration of marine viruses in culture and natural samples by flow cytometry. *Appl. Environ. Microbiol.* 65, 45–52.
- Martin-Cuadrado, A. B., Lopez-Garcia, P., Alba, J. C., Moreira, D., Monticelli, L., Strittmatter, A., et al. (2007). Metagenomics of the deep mediterranean, a warm bathypelagic habitat. *PLoS One* 2:e914. doi: 10.1371/journal.pone.0000914
- Martinez Martinez, J., Swan, B. K., and Wilson, W. H. (2014). Marine viruses, a genetic reservoir revealed by targeted viromics. *ISME J.* 8, 1079–1088. doi: 10.1038/ismej.2013.214
- Martinez-Hernandez, F., Fornas, O., Lluésma Gomez, M., Bolduc, B., De La Cruz Pena, M. J., Martinez, J. M., et al. (2017). Single-virus genomics reveals hidden cosmopolitan and abundant viruses. *Nat. Commun.* 8:15892. doi: 10.1038/ncomms15892
- McGillicuddy, D. J. Jr., Anderson, L. A., Bates, N. R., Bibby, T., Buesseler, K. O., Carlson, C. A., et al. (2007). Eddy/wind interactions stimulate extraordinary mid-ocean plankton blooms. *Science* 316, 1021–1026. doi: 10.1126/science.1136256
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Goker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. doi: 10.1186/1471-2105-14-60
- Meier-Kolthoff, J. P., and Goker, M. (2017). VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* 33, 3396–3404. doi: 10.1093/bioinformatics/btx440
- Meier-Kolthoff, J. P., Hahnke, R. L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., et al. (2014). Complete genome sequence of DSM 30083(T), the type strain (U5/41(T)) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci.* 9:2. doi: 10.1186/1944-3277-9-2
- Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257. doi: 10.1038/ncomms11257
- Middelboe, M., Jorgensen, N., and Kroer, N. (1996). Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. *Appl. Environ. Microbiol.* 62, 1991–1997.
- Middelboe, M., and Lyck, P. G. (2002). Regeneration of dissolved organic matter by viral lysis in marine microbial communities. *Aqua. Microbial. Ecol.* 27, 187–194. doi: 10.3354/ame027187
- Mizuno, C. M., Ghai, R., Saghai, A., López-García, P., and Rodríguez-Valera, F. (2016). Genomes of abundant and widespread viruses from the deep ocean. *mBio* 7, e805–e816. doi: 10.1128/mBio.00805-16
- Nagata, T., Fukuda, H., Fukuda, R., and Koike, I. (2000). Bacterioplankton distribution and production in deep Pacific waters: large-scale geographic variations and possible coupling with sinking particle fluxes. *Limnol. Oceanogr.* 45, 426–435. doi: 10.4319/lo.2000.45.2.0426
- Nigro, O. D., Jungbluth, S. P., Lin, H.-T., Hsieh, C.-C., Miranda, J. A., Schwarcz, C. R., et al. (2017). Viruses in the oceanic basement. *mBio* 8, e2129–e2116. doi: 10.1128/mBio.02129-16
- Nunoura, T., Takaki, Y., Hirai, M., Shimamura, S., Makabe, A., Koide, O., et al. (2015). Hadal biosphere: insight into the microbial ecosystem in the deepest ocean on Earth. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1230–E1236. doi: 10.1073/pnas.1421816112
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Simpson, J., and Solymos, P. (2018). *Vegan: Community Ecology Package*. Available at: <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- Parada, V., Sintes, E., Van Aken, H. M., Weinbauer, M. G., and Herndl, G. J. (2007). Viral abundance, decay, and diversity in the meso- and bathypelagic waters of the north atlantic. *Appl. Environ. Microbiol.* 73, 4429–4438. doi: 10.1128/aem.00029-07
- Rambaut, A. (2006). *FigTree 1.4.3 - A Graphical Viewer of Phylogenetic Trees and a Program for Producing Publication-Ready Figures*. Available at: <https://mac.softpedia.com/get/Graphics/FigTree.shtml>, <https://www.ed.ac.uk/edinburgh-infectious-diseases/research/themes/disease-dynamics/andrew-rambaut>
- Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212. doi: 10.1038/nature08060
- Roux, S., Enault, F., Hurwitz, B. L., and Sullivan, M. B. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3:e985. doi: 10.7717/peerj.985
- Schmittner, A., Chiang, C. H., and Hemming, R. S. (2007). “The ocean’s meridional overturning circulation,” in *Ocean Circulation: Mechanisms and Impacts – Past and Future Changes of Meridional Overturning*, eds A. Schmittner, C. H. Chiang, and S. R. Hemming (Washington, DC: American Geophysical Union).
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sintes, E., Bergauer, K., De Corte, D., Yokokawa, T., and Herndl, G. J. (2013). Archaeal amoA gene diversity points to distinct biogeography of ammonia-oxidizing Crenarchaeota in the ocean. *Environ. Microbiol.* 15, 1647–1658. doi: 10.1111/j.1462-2920.2012.02801.x
- Sintes, E., De Corte, D., Ouillon, N., and Herndl, G. J. (2015). Macroecological patterns of archaeal ammonia oxidizers in the atlantic ocean. *Mol. Ecol.* 24, 4931–4942. doi: 10.1111/mec.13365
- Sul, W. J., Oliver, T. A., Ducklow, H. W., Amaral-Zettler, L. A., and Sogin, M. L. (2013). Marine bacteria exhibit a bipolar distribution. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2342–2347. doi: 10.1073/pnas.1212424110
- Talley, L. D. (2013). Closure of the global overturning circulation through the indian, pacific, and southern oceans schematics and transports. *Oceanography* 26, 80–97. doi: 10.5670/oceanog.2013.07
- Wigington, C. H., Sonderegger, D., Brussaard, C. P., Buchan, A., Finke, J. F., Fuhrman, J. A., et al. (2016). Re-examination of the relationship between marine virus and microbial cell abundances. *Nat. Microbiol.* 1:15024. doi: 10.1038/nmicrobiol.2015.24
- Wommack, K. E., Hill, R. T., Muller, T. A., and Colwell, R. R. (1996). Effects of sunlight on bacteriophage viability and structure. *Appl. Environ. Microbiol.* 62, 1336–1341.
- Woyke, T., Sczyrba, A., Lee, J., Rinke, C., Tighe, D., Clingenpeel, S., et al. (2011). Decontamination of MDA reagents for single cell whole genome amplification. *PLoS One* 6:e26161. doi: 10.1371/journal.pone.0026161

- Yang, Y., Yokokawa, T., Motegi, C., and Nagata, T. (2014). Large-scale distribution of viruses in deep waters of the Pacific and Southern Oceans. *Aqua. Microbial Ecol.* 71, 193–202. doi: 10.3354/ame01677
- Yokokawa, T., Yang, Y., Motegi, C., and Nagata, T. (2013). Large-scale geographical variation in prokaryotic abundance and production in meso- and bathypelagic zones of the central Pacific and Southern Ocean. *Limnol. Oceanogr.* 58, 61–73. doi: 10.4319/lo.2013.58.1.0061
- Zeng, Q., and Chisholm, S. W. (2012). Marine viruses exploit their host's two-component regulatory system in response to resource limitation. *Curr. Biol.* 22, 124–128. doi: 10.1016/j.cub.2011.11.055

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 De Corte, Martínez Martínez, Cretoiu, Takaki, Nunoura, Sintes, Herndl and Yokokawa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.